

# Chancen im OPUS:

## Automatisiert SWD-Schlagwörter produzieren

---

Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Inhalterschließung 2009: Potentiale nutzen  
33. Jahrestagung der Gesellschaft für Klassifikation  
Technische Universität Dresden  
18. März 2009



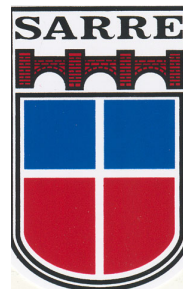
## SULB: Aufgabenspektrum

---

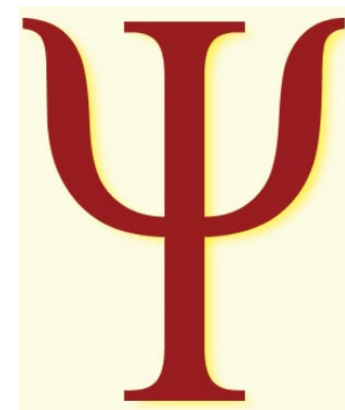
Saarländische Universitäts- und Landesbibliothek  
SULB



Universitätsbibliothek der  
Universität des Saarlandes



Landesbibliothek des Saarlandes



Sondersammelgebiet  
Psychologie

---

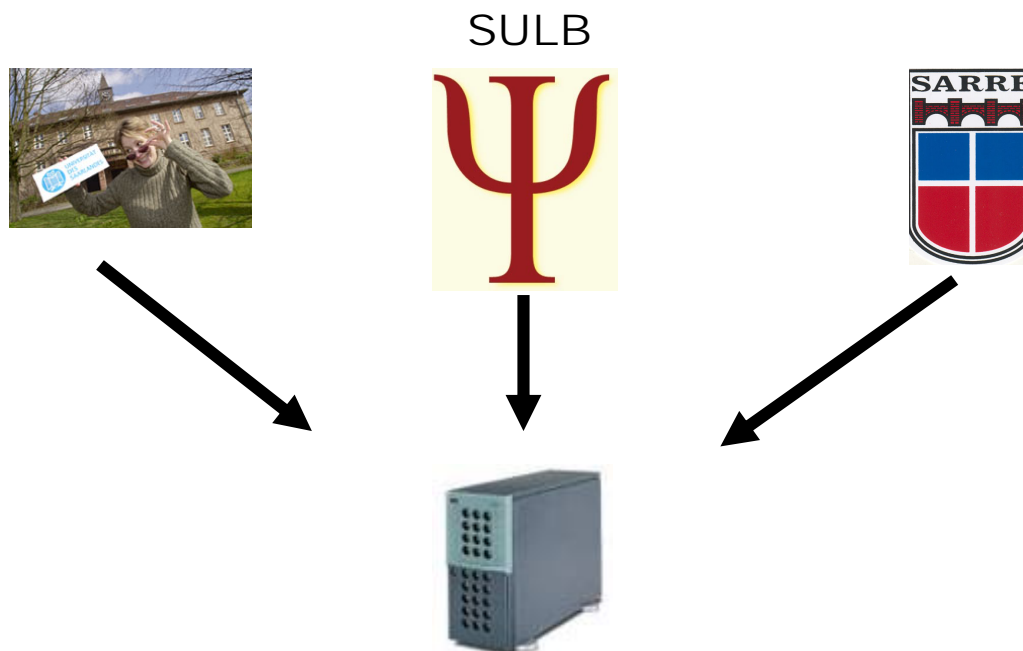
Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Inhalterschließung 2009: Potentiale nutzen  
33. Jahrestagung der Gesellschaft für Klassifikation  
Technische Universität Dresden  
18. März 2009



## SULB: ein Repository für alle Aufgaben

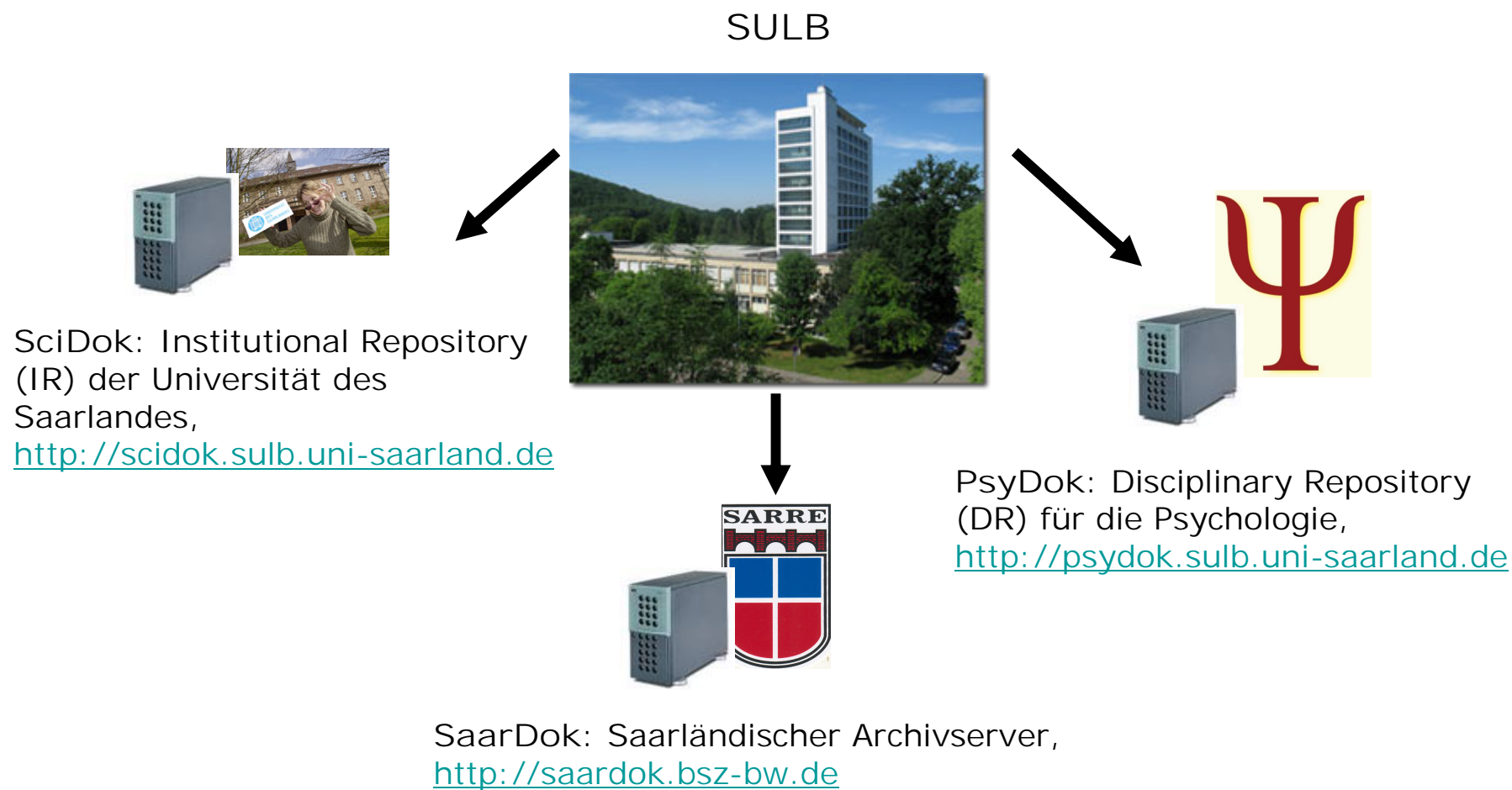
---



Sova: Saarbrücker Online Volltext Archiv

.... bis 2003

# SULB: je Aufgabe ein Repository



## Repositories: Technik

---

- SciDok und PsyDok  
Software: OPUS, DINI-zertifiziert  
-> PUSH-Service: Wissenschaftler spielen Dokumente ein
- SaarDok  
Software: anfangs OPUS-basiert  
später eine Instanz des BSZ-Online-Archivs BOA,  
jetzt: SWBContent, gehostet vom BSZ  
-> PULL-Service: Bibliothekare harvesten Dokumente



## Workflow in Repositories

---

... in typischen Repository-Umgebungen (wie SciDok und PsyDok)

- Autor füllt Erfassungsmaske aus und erstellt dabei Metadaten
- Autor lädt Dokument zu den Metadaten auf den Server
- Metadaten und Dokument liegen im temporären Bereich des Repositorys
- Bibliothekare überprüfen Dokument und Metadaten
- Bibliothekare schalten nach erfolgreicher Kontrolle das Dokument frei



## Metadaten in Repositories

---

### Typische Metadaten-Felder:

- Titel
- Verfasser/in
- Institut
- Urheber
- Beteiligte Institutionen
- Freie Schlagworte
- Abstract
- Jahr der Fertigstellung
- Quelle: Zeitschrift, Sammelband, alternative Publikation

...

### Sacherschließung:

- Fachklassifikation
- Schlagwortnormdatei SWD



# Repositories & SWD: das Problem

---



<http://commons.wikimedia.org/wiki/File:Scientist.svg>





## Probleme bei SWD-Vergabe

---

Leider ist den Autoren die SWD und ihre Nutzung unvertraut

Konsequenz a)

Autoren können das Dokument zwar sehr exakt beschreiben,  
aber ...

- sie vergeben unscharfe und zu allgemeine Schlagworte
- sie vergeben schlimmstenfalls falsche Schlagworte
- oder sie ignorieren die SWD und verwenden Begriffe, die in der SWD nicht abgebildet sind



## Metadaten

---

### Konsequenz b)

Bibliothekare verfügen über die notwendige SWD-Expertise und müssen die vorgenommene Verschlagwortung aufwändig nachbearbeiten oder – ohne Vorschläge durch Autoren – komplett selbst vornehmen, ...

können aber die Dokumente inhaltlich nicht so exakt beschreiben wie die Autoren.



## Verallgemeinerbarkeit?

---

Sommer 2007, Frühjahr 2009: SULB erhebt Status Quo unter Repository-Betreibern (OPUS-, DSpace-, EPrints- und MyCoRe-Systeme sowie Benutzer von Eigenentwicklungen)

Fazit:

Größtenteils analoge Erfahrungen ...

- Autoren vergeben falsche/unpassende Schlagworte
- (Nach-)Bearbeitung durch eigene Mitarbeiter
- Vergabe von SWD-Schlagworten im Repository wurde wegen Problemen/Aufwand eingestellt

...



## Warum SWD in Repositories?

---

- Sie garantiert die Erschließung aller Dokumente über ein fachübergreifendes, kontrolliertes Vokabular.
  - Sie sichert die nahtlose Integration der Metadaten in bibliotheksspezifische Retrievalsysteme und Anwendungen.
  - Auf aggregierter Ebene können Dokumente systematisch und unter Verwendung eines gepflegten und strukturierten Schemas recherchiert werden.
- > Vergabe von SWD-Schlagworten ist eine Empfehlung im Zertifikat für Dokumenten- und Publikationsservices der Deutschen Initiative für Netzwerkinformation e.V. (DINI)
- > Zahlreiche Repositories verlangen Vergabe von SWD-Schlagworten



## SWD & Repositories: Ein Entwurf

---

Technische Unterstützung der Autoren bei der Auswahl treffender SWD-Schlagworte, um

- a) über eine exakte Verschlagwortung ein optimiertes Retrieval zu ermöglichen und
- b) eine Erleichterung des Veröffentlichungsprozesses (easy submission) und
- c) Bearbeitungsprozesses zu erreichen.

Demo: <http://psydok.sulb.uni-saarland.de/demo/bibtag2008/>



## Was passiert?

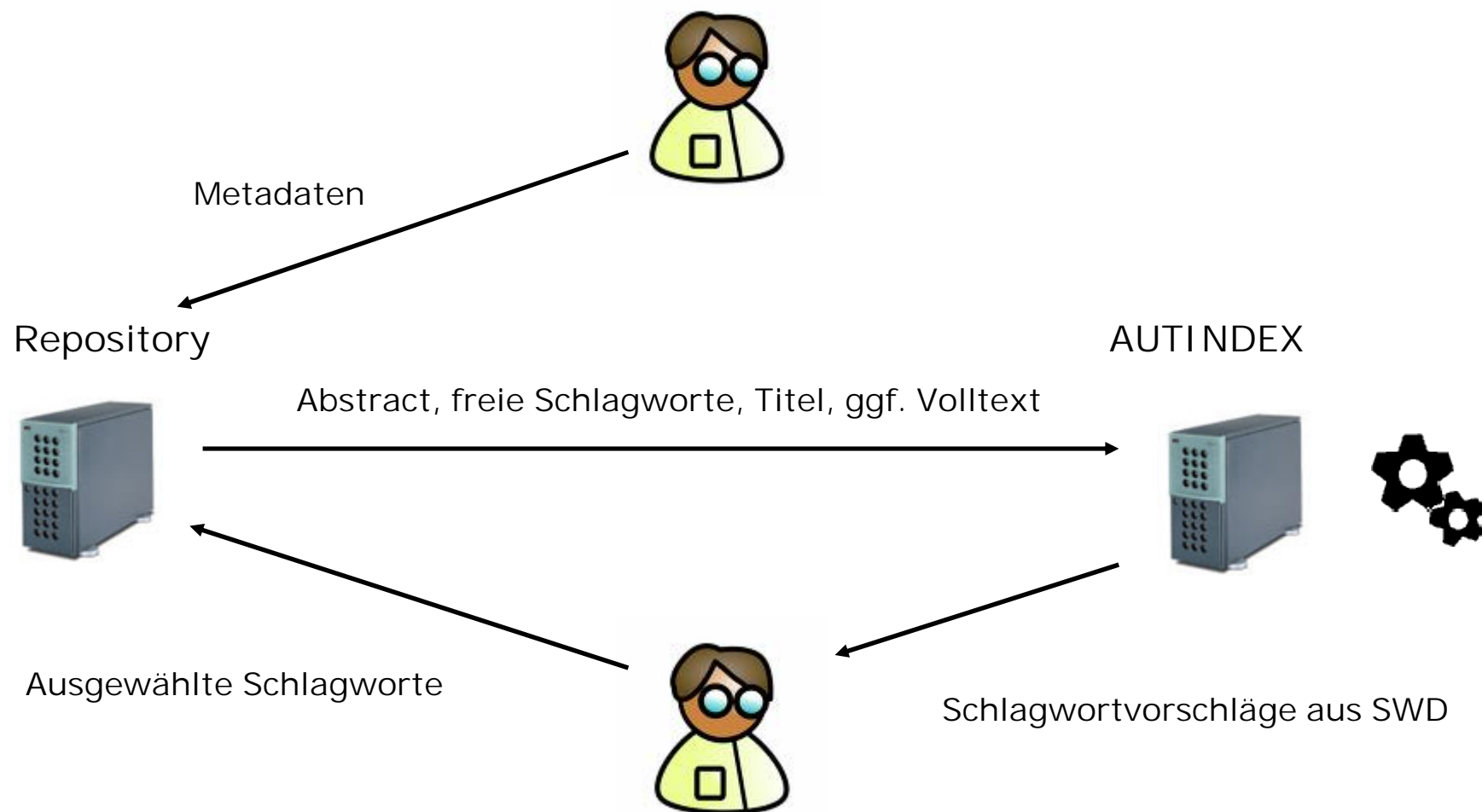
---



<http://commons.wikimedia.org/wiki/File:Scientist.svg>

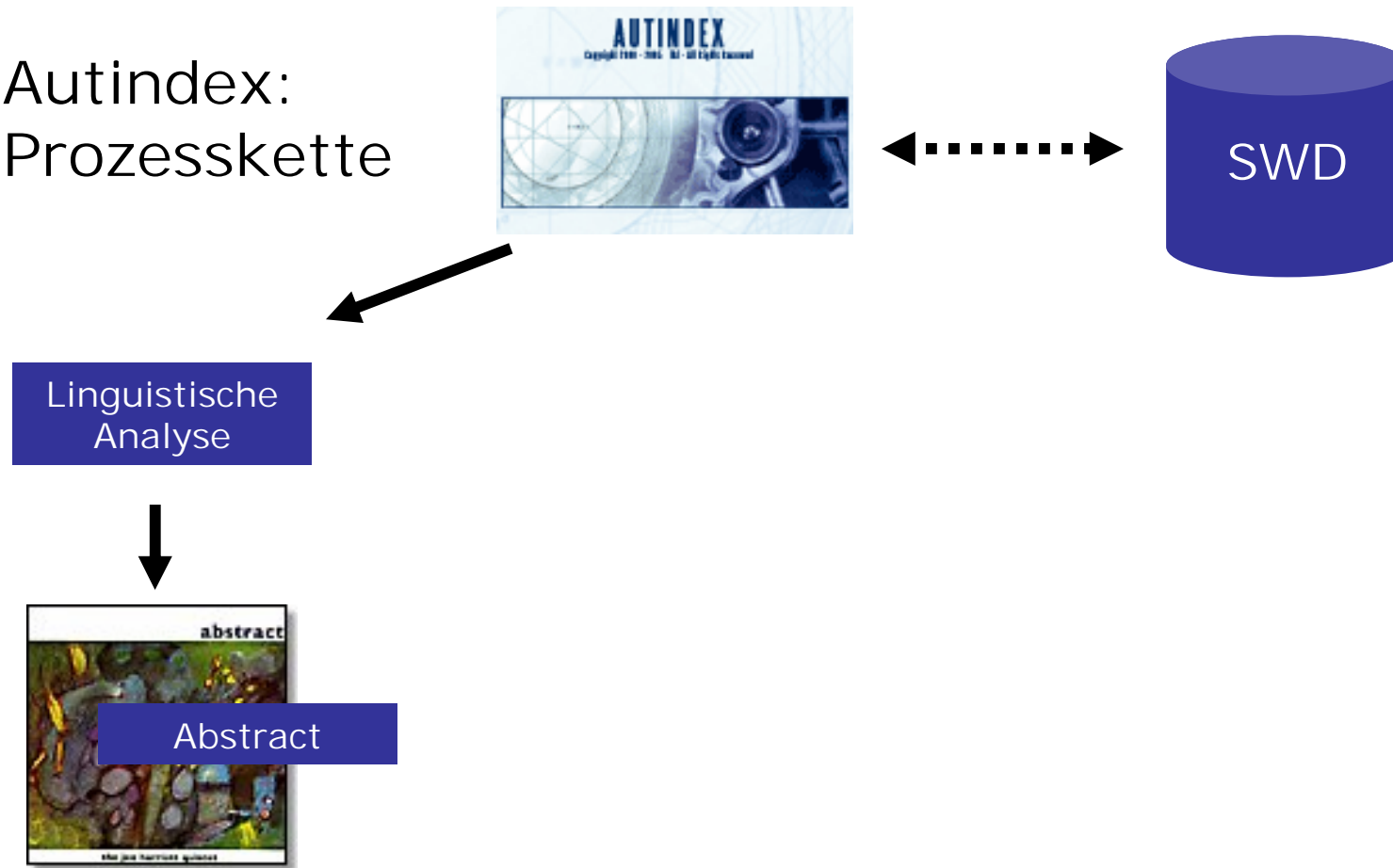


## Was passiert?



## Funktionsweise: Autindex

### Autindex: Prozesskette





## Funktionsweise: Linguistisches Verfahren

---

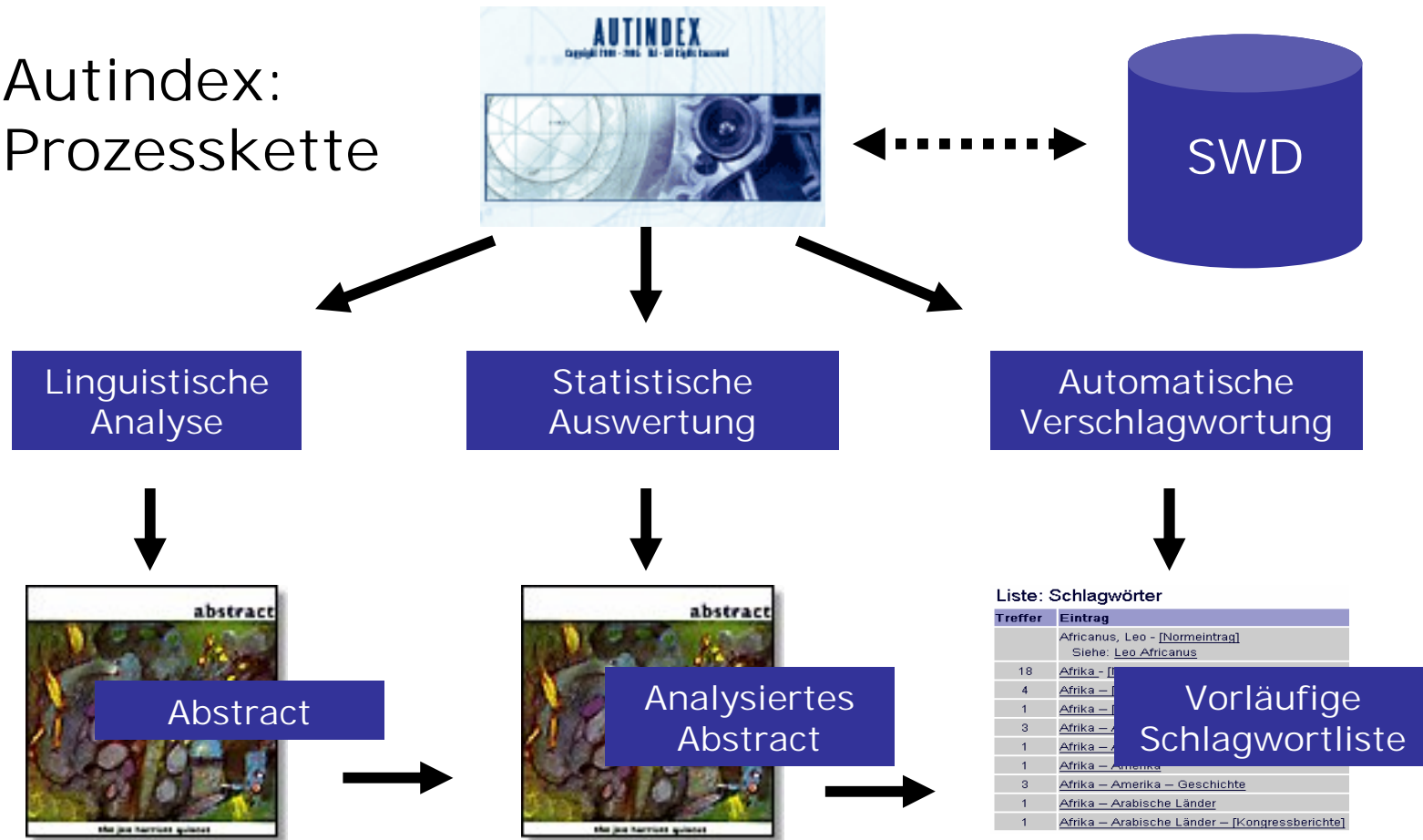
### Autindex: Vorteile des linguistischen Verfahrens bei der Textanalyse

- Erkennung morphologischer Varianten (Wittgenstein / wittgensteinsche Sprachphilosophie, Haus / Häusern)
- Kompositazerlegung: Bessere Information über Semantik des Texts: Rentenversicherungsangestellter à Rente, Versicherung, Angestellter / Rentenversicherung, Versicherungsangestellter
- Erkennung syntaktischer Varianten: (Reduktion von/der Kosten / Kostenreduktion, Reduzierung von Kosten, Kostenreduzierung)
- Semantische Varianten (Synonyme: Müllwirtschaft / Abfallwirtschaft)
- Kombination mit syntaktischen Varianten: (Kostenverminderung / Verminderung von Kosten / Einsparungen der Kosten)



# Funktionsweise: Autindex

## Autindex: Prozesskette



---

## Geplante Realisierung:

Im Rahmen eines gemeinsamen bei der Deutschen Forschungsgemeinschaft DFG beantragten Projekts des Institutes der Gesellschaft zur Förderung der Angewandten Informationsforschung e.V. an der Universität des Saarlandes ([IAI](#)) und der Saarländischen Universitäts- und Landesbibliothek ([SULB](#))



## Kontaktinformationen

---

Vielen Dank für Ihre Aufmerksamkeit!

Ulrich Herb

Saarländische Universitäts- und Landesbibliothek  
Gebäude B1 1, Zi. 9.08,  
D-66123 Saarbrücken  
Telefon: 0049 681 302-2798  
[u.herb@sulb.uni-saarland.de](mailto:u.herb@sulb.uni-saarland.de)

---

Ulrich Herb,  
Saarländische Universitäts- und  
Landesbibliothek SULB

Inhaltserschließung 2009: Potentiale nutzen  
33. Jahrestagung der Gesellschaft für Klassifikation  
Technische Universität Dresden  
18. März 2009

